



**Figure S2:** Empirical distributions of Tajima's  $D$  and  $\pi$  built using the 316 genes surveyed by SeattleSNPs (<http://pga.gs.washington.edu/>). Genes within the upper extreme (97<sup>th</sup>) of the distribution are marked in grey. Gene classes with values close to survey genes are indicated by *SERPINB11* gene name. **A - C:** YRI sample – SeattleSNPs panel 2; **D - F:** All African descent samples – SeattleSNPs panels 1 to 3. *ABO*: ABO blood group; *CD151*: CD151 antigen; *CYP4F3*: Cytochrome P450, family 4, subfamily F, polypeptide 3; *FLG1*: Fibrinogen-like 1; *FUT2*: Fucosyltransferase 2; *GPR109G*: G protein-coupled receptor 109B.